12-2009

Rigour versus Relevance revisited: Evidence from IS Conference Reviewing Practice

Michael Rosemann
Information Systems Program, Queensland University of Technology, m.rosemann@qut.edu.au

Jan Recker
Information Systems Program, Queensland University of Technology, j.recker@qut.edu.au

Rigour versus Relevance Revisited: Evidence from IS Conference Reviewing Practice

Michael Rosemann and Jan Recker
Information Systems Program
Queensland University of Technology
Brisbane, Australia
Email: {m.rosemann, j.recker}@qut.edu.au

Abstract

A commonly held belief in the IS discipline is that rigour and relevance are contrary to each other and tha addressing both is virtually impossible. It is also believed widely that the editorial practices of our premie conferences and journals over-emphasise rigour on the cost of relevance. However, while these two topics have been filled with numerous subjective discussions, more solid evidence into the true relationship between rigou and relevance and the impact of conference editors on this relationship is still outstanding. This pape contributes to this debate by deriving empirical evidence from a comprehensive and detailed analysis of the characteristics of the submissions and the reviewing practices of three recent IS conferences. It provides firs insights into the actual relationship between rigour and relevance and into the role conference chairs play i balancing rigour and relevance. Besides the outcomes that the current set of evaluation criteria does not provide a straight forward proxy for relevance to practitioners, the paper offers two main contributions. First, empirica insights are provided that rigour and relevance do in fact not have to be mutually exclusive. Second, the editoria practices at conferences are skewed towards rigorous papers rather than relevant papers.

Keywords

Relevance, rigour, academic research.

INTRODUCTION

In their MISQ commentary in 1999 that (re-)sparked the rigour versus relevance debate, Benbasat and Zmud (1999) argued that the IS discipline overemphasises rigour on cost of relevance. They put the case forward that the legitimacy of our discipline may be at risk if IS research fails to produce relevant outcomes that will educate the next wave of IS professionals, as well as provide fertile ground for disseminating knowledge and providing leadership to organizations on the effective management and utilization of information technologies. Accordingly, Benbasat and Zmud (1999) suggested a set of guidelines – both for authors and reviewers of scholarly work – on how to improve, practice and demonstrate relevance in IS articles without compromising rigour. Eight of their recommendations relate to how authors can improve their work, and have, over time, motivated other researchers to address these recommendations (e.g., Davenport and Markus 1999; Rosemann and Vessey 2008). Notwithstanding the impact that these author guidelines may have had, very little – if any – progress has been made in addressing Benbasat and Zmud's (1999, p. 12) ninth recommendation:

"Editors and editorial boards of all IS journals need to critically examine their current postures, reviewing procedures, and editorial decisions concerning the balance between rigour and relevance with the goal of publishing manuscripts that are characterized by both."

Benbasat and Zmud (1999) clearly argue that editors and reviewers should strive to look for a balance of rigorous research conduct *and* relevant research in manuscripts they are committed to review. Clearly, such a recommendation addresses journal reviewing practice as well as conference reviewing practice. And indeed, in the study reported in this paper, we consider IS conference practices over IS journal practices. We have done so for three main reasons. First, the deadline-driven submission process of a conference allows a comprehensive comparative analysis of a large number of submissions that are all evaluated based on the same set of criteria, within a short timeframe. Second, little research of a similar nature has been conducted on conference papers; for example, the majority of bibliometric studies has focused on journal articles (e.g., Baskerville and Myers 2002). Third, there is growing evidence that journal publications are not necessarily representative of the IS field as a whole (Avgerou, Siemer-Matravers and Bjørn-Anderson 1999; Whitley and Galliers 2007).

We are interested in a set of questions fundamental to the legitimacy of the IS discipline: Does our discipline adhere to rigour and relevance recommendations such as the ones voiced by Benbasat and Zmud (1999)? Is relevance a significant factor in acceptance/rejection decisions at IS conferences? If so, does it count as much as

rigour? And ultimately, what kind of articles are being published by IS conferences – rigorous ones, relevant ones, or the ones that are both rigorous and relevant?

In order to be able to meaningfully explore such questions, however, we believe it is worthwhile to study if the set of papers submitted to IS conferences might already be biased in terms of the balance between rigour and relevance. In other words: Could it be that in fact not the conference editors, but the authors of the submitted papers bring the rigour-relevance relationship out of balance by submitting papers that overemphasize rigour on cost of relevance (or vice versa)? Such a situation would, for instance, be reflected in significantly different scores for the quality of the rigour of a paper as opposed to its perceived relevance.

However, besides a "hunch" that this may be the cause, and aside from a plethora of subjective views on this topic, there is no evidence to indicate that authors indeed have a tendency or motivation to rate one (rigour or relevance) higher than the other. Accordingly, we tentatively formulate the following research proposition that we seek to explore in this paper:

P1: Based on the pool of papers submitted to a conference, there is no indication that rigour and relevance are contradictory to each other.

A confirmation of this proposition would mean that there is *no* compromise required in the relationship between rigour and relevance because conference submissions are successful in meeting required standards both for rigour and relevance. This could be visualised in a scatter plot such as the one shown on the left hand side in Figure 1. In a diagram that is based on data from a conference review system, and that plots all submitted papers on the axes *perceived quality of the rigour* of the paper and *perceived quality of the relevance* could look like the shape on the left, which, in turn, would confirm proposition P1. The sum of all papers would form the depicted shape and could be approximated by a regression analysis that leads to a linear curve with a 45 degree angle. In this case a higher ranked paper dominates a lower ranked paper in both rigour and relevance.

Assuming our research proposition P1 would not hold, we would achieve a scatter plot shaped like the one depicted in the diagram on the right side of Figure 1. In this case, conference submissions are characterized by either a high degree of perceived rigour (or relevance), or a low score on both scales. Searching for the set of papers to be accepted in such a case, there is probably no opportunity to accept papers that indeed balance rigour and relevance – and achieve high scores on both scales. In fact, a high level of rigour most likely negatively correlates with the level of relevance, and vice versa. Such a shape could be approximated by an isoquant. Such a shape would indicate the existence of a rigour *versus* relevance relationship.

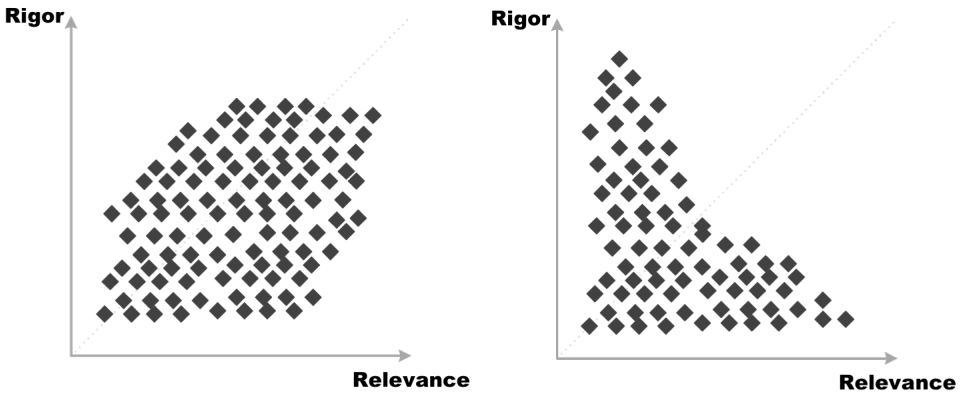

Figure 1: Two extreme forms of the rigour-relevance relationship: complementary (left) versus contradictory (right)

Assuming we are able to obtain empirical insights into the characteristics of IS conference submissions, and to establish the extent to which these submissions are characterized by either a correlated or contradictory relationship between rigour and relevance, the next question would then be which of these submissions ultimately get accepted at IS conferences – the ones that score on the rigour scale, high on the relevance scale, or the ones that achieve both?

Accordingly, our second research proposition is related to the impact of conference editors on the rigour-relevance relationship. As conference editors have the opportunity to weight the evaluation criteria of a conference review system differently, explicitly or implicitly, they have impact on the importance of these characteristics to the ultimate acceptance decision. Also, conference editors make the ultimate decision to accept or reject a paper, based on the reviews received but also taking into account other criteria (such as the meeting a desired acceptance rate, papers allowed per track, proceedings paper limits and so forth). Following the general tenet of the ongoing rigour vs. relevance debate in IS (e.g., Benbasat and Zmud 1999; Agarwal and Lucas Jr. 2005) that IS research overemphasises rigour on cost of relevance (e.g., Davenport and Markus 1999; Recker et

al. 2009), we believe that there are strong signs suggesting that the IS community puts more emphasis on rigour than relevance. Thus, we tentatively state our second research proposition:

P2: Conference editors rank rigour higher than the relevance when making the acceptance/rejection decision.

A confirmation of this proposition would be the first empirical evidence that indeed the rigour of a conference paper is more important than its relevance to the acceptance decision. In this case, the editors of conferences could potentially be substantial gate keepers in the process of disseminating IS research outcomes, or even research impact to practice. In order to visualise the alternative impact of conference editors on the rigour-relevance relationship, we depicted three potential alternative cases in Figure 2. The point on the bottom left is the coordinate that represents the average score (and standard deviation) for the quality of the rigour and the quality of the relevance across all submitted papers. The impact of the editors can now be seen by the way this point is moved in the rigour-relevance diagram when only the accepted papers are selected, and a coordinate reflecting the average score of rigour and relevance of these accepted papers is depicted. Scenario I shows the case that the decision of the editors is skewed towards a higher emphasis on rigor rigour – which is congruent to our expectation expressed in research proposition P2. Scenario II moves the initial point along a 45 degree curve and represents an equal emphasis on rigour and relevance, with accepted papers scoring higher (in comparison to all papers) on both criteria. Scenario III, finally, shows the scenario that the editors put more emphasis on the relevance of a paper than on its scientific.

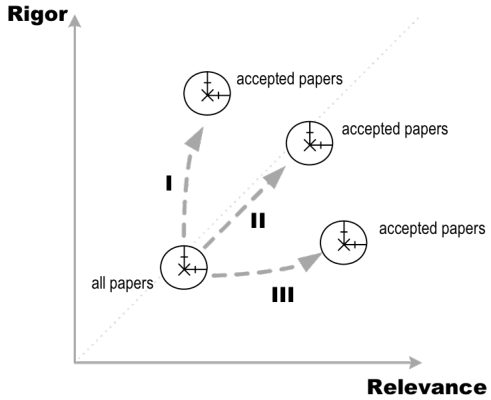

Figure 2: Alternative scenarios for the possible impact of conference editors on paper acceptance

The aim of our research is to test these two propositions. We collected review data from three IS conferences and analysed the data. Our objective in writing this paper is to seek empirical insights, where possible, into the actual relationship between rigour and relevance on two levels:

(1) the pool of papers submitted to an IS conference differentiated,

(2) the impact of the decision of conference editors in terms of the relative importance of these criteria regarding the acceptance/rejection decisions.

We proceed as follows. The next section revisits the rigour versus relevance debate in our discipline. Then, we introduce the conferences from which we obtained the review data. Next, we present and discuss our analysis of the review data. Following that, we discuss first empirical insights on the relationship between rigour and relevance and the impact of the editors in conference reviewing practice. We then conclude this paper by reviewing our contributions and suggesting guidelines for both IS authors and reviewers.

RIGOUR VERSUS RELEVANCE: THE DEBATE

The debate on rigour versus relevance (e.g., Benbasat and Zmud 1999; Agarwal and Lucas Jr. 2005) in IS research has been raised frequently within the IS community over the last decade. Commentaries have appeared in MIS Quarterly (Vol. 23, Issue 1, 1999) and as part of specific research methodologies (for example, MIS Quarterly, Vol. 28, Issue 3, 2004, with its focus on Action Research). Attention has also been paid to each of the two sides of the debate in its own right; in terms of research relevance, for instance, in a special issue of the Information Resources Management Journal (e.g., Robey and Markus 1998; Senn 1998), and as a number of papers in Communications of the AIS (Vol. 6, 2001). Similarly, research rigour has been the focus of a substantial number of papers and special issues (e.g., Grover, Lee and Durand 1993; Dubé and Paré 2003). Likewise, IS conferences have taken on, and contributed to, the rigour versus relevance debate. The 2007 European Conference on Information Systems (ECIS), for instance, was subtitled "Relevant Rigour – Rigorous Relevance." Similarly, the 2007 Australasian Conference on Information Systems was titled "The 3Rs: Research, Relevance, Rigour – Coming of Age", and featured a panel debate on the R versus R debate (Recker et al. 2009).

The main argument of the debate typically goes like this: For IS research to establish itself as a discipline in its own right, relevance must not be lost in the face of rigorous research conduct. At the same time, while being relevant, rigour must be applied to create reliable, transparent and valid research outcomes. In this context, rigour is defined as the extent to which scholarly work is of satisfactory research quality, uses a sound methodology and meets commonly accepted reliability and validity standards (Benbasat and Weber 1996; Rosemann and Vessey 2008). In contrast, relevance is the extent to which scholarly work is interesting and important, applicable, current and accessible to practitioners (Benbasat and Zmud 1999; Klein, Jiang and Saunders 2006).

However, notwithstanding the noble objective of fostering and encouraging papers to be both relevant and rigorous (Benbasat and Weber 1996), it is not at all clear that the IS academic community has made much progress in recent years. The issue of rigour, or lack thereof, for instance, in published IS articles has often been lamented, and a number of guidelines have been proposed in the past to increase the rigour of various research methods (e.g., Dubé and Paré 2003; Straub, Boudreau and Gefen 2004). The strong emphasis on rigour supports the prevalent common perception that IS academics need to conduct rigorous research to guarantee that it is of high quality in order to establish credibility, to publish in high quality journals, to attain tenure and promotion, and to compete for research funding (Robey and Markus 1998; Dennis et al. 2006).

Following this strong emphasis on rigour, many authors argue that the focus on rigour leads to a negligence of relevance in IS publications (e.g., Robey and Markus 1998; Benbasat and Zmud 1999). Rosemann and Vessey (2008), for instance, identify two major issues that frequently obstruct IS scholars from placing more value on relevance. First, there are few incentives for researchers to conduct research that is relevant to practice. This is most likely due to the perception that it is not valued by the top IS research journals and conferences, and neither by promotion committees. And indeed, the fact that IS publications outlets that emphasise relevance over rigour, have, of late, been paying a harsh price, for instance, by suffering in journal rankings (consider, for example, Communications of the ACM or Journal of Information Technology). Second, there is the perception that it is simply impossible to attain both rigour and relevance in research.

These and other statements (e.g., Davenport and Markus 1999; Recker et al. 2009) suggest that the IS academic community most often views rigour and relevance as conflicting research objectives, that is, as the two extreme points of a continuum. Achieving one is viewed as necessarily compromising the other. In other words, the issue is perceived as that of rigour *versus* relevance. This view could explain why a lot of papers published at IS venues are often viewed as not relevant (Darroch and Toleman 2005) – because relevance may be seen as a paper characteristic that has to be compromised so that required rigour standards of conferences and/or journals can be met. This conflicting view between rigour and relevance is also present in the way some – or most – of the typical IS funding bodies distribute their resources: Funding by industry groups or commercial companies tends to overemphasize relevance without regard to rigour. Funding by government research agencies tends to overemphasize rigour without regard to relevance. This makes it hard to believe that the institutional bodies that govern our research actually have an interest in a middle ground between relevance and rigour (Baskerville 2007).

While these arguments suggest a conflicting relationship between rigour and relevance in IS publications, other authors argue that it could be, or at least should be, possible to conduct research that is both rigorous *and* relevant (Benbasat and Weber 1996; Robey and Markus 1998; Senn 1998). Accordingly, while the passion for this topic is without any doubt high, the guidelines offered undisputedly valuable, and more clarity indispensable for shaping the future IS discipline, we assert that, still, little factual knowledge exists regarding the *actual* relationship between rigour and relevance.

The fact that, over time, a wide range of guidelines has been suggested to IS scholars to improve rigour and relevance of manuscript submissions, would suggest that papers submitted to IS conferences and journals have improved in terms of relevance and rigour. Whether this is actually the case, however, remains to be shown. Accordingly, our objective in writing this paper is to provide initial empirical evidence about the actual relationship between rigour and relevance based on an examination of the reviewing practices of three IS conferences.

DATA BACKGROUND: THE SELECTED CONFERENCES

Background to the Conferences

We examined the reviewing practices of three IS conferences, viz., the 15th European Conference on Information Systems (ECIS) 2007; the 5th International Business Process Management Conference (BPM) 2007; and the 26th International Conference on Conceptual Modeling (ER) 2007.

We selected ECIS due to its standing as one of the world's top three IS conferences, and the fact that it is the largest and most prestigious European IS conference (Whitley and Galliers 2007). ECIS 2007 featured 16 different tracks across a wide range of IS research domains, including IS research methodologies, organizational

engineering, e-work, IS security, IS economics, knowledge management, and others. Overall, ECIS 2007 accepted 200 papers from 580 submissions, representing an acceptance rate of 34.5 per cent.

We selected the BPM conference because business process management and the development and use of process-aware information systems is an emerging IS research domain that is characterized by high relevance to current business and management practice (Rosemann 2008). Major IS conferences (e.g., ACIS 2007, AMCIS 2007, ECIS 2008, HICSS 2008-2010) feature dedicated tracks on business process management in their conference program. BPM 2007 received in total 152 submissions, of which 21 were accepted, an acceptance rate of 13.8 per cent.

The ER conference is the most reputable and competitive conference on one of the core research themes in IS, conceptual modelling for IS analysis and design (Wand and Weber 2002). ER provides the most prestigious annual forum for exploring research, development, novel applications, and industrial innovations in the area of conceptual modelling and associated phenomena. ER 2007 received 160 full paper submissions, 38 of which were accepted, an acceptance rate of 23.8 per cent.

Data Collection

To gain insights into the review practices of the three conferences, we proceeded as follows: First, we approached key members of the respective Conference Organizing Committees, to inform them about intent and purpose of our study. Next, upon agreement, we requested review data in anonymous format for the quantitative evaluations of each of the papers. Specifically, to ensure anonymity, we asked not to be given any identifying data about the manuscripts (e.g., submission IDs, paper titles, author names, qualitative review comments), nor any identifying information about the review or editorial teams (e.g., reviewer ID, reviewer names, associate editors, track chairs or program committees).

In information systems, each conference establishes its own review criteria. A number of types of data typically are used. First, each of the conferences requested reviewers to rate papers using well-defined quantitative criteria. Second, the reviewers provide an overall evaluation score as a suggestion to the program committee members and track chairs in aiding the ultimate acceptance/rejection decision. Third, program committee members and track chairs consider the qualitative, written reviews, in addition to the quantitative paper review scores. These written comments are made by the reviewers in support of their decision, and also to give input to the paper authors as to how the paper could be improved. Fourth, other factors such as the number of submissions per track, and so on, may influence the final acceptance decision made by the conference editors.

While these four types of data are all instrumental in the reviewing process, we focus on the effect of the scores on the objective review criteria and the overall evaluation of the paper by the reviewers on the acceptance/rejection decision made by the editors. We believe this focus to be appropriate given that program committees are often required to evaluate hundreds if not thousands of submissions in a short timeframe. Scores on review criteria and the overall evaluation are therefore a fundamental source of information in the decision process. At the same time, such analyses allow us to determine the extent to which subjective considerations come into play in the decision.

We examined the review systems of the three conferences to determine similarities and differences in their reviewing practices. The ECIS review system presented a brief description of the evaluation criteria, while the other two review systems simply listed the evaluation criteria without further clarification. Some criteria were similar, if not common, across all conferences (such as *relevance to the conference theme* and *presentation*). Further, all review systems requested reviewers to rate the submissions using an *overall evaluation* score. However, we also found also some differences. For example, the review system of the BPM conference captured the *perceived confidence* of the reviewer in their judgment. The confidence score is used to weight the scores for the individual criteria. Similar to the BPM conference, the ER review system captured the *perceived expertise* of the reviewer in order to classify the reviewer's confidence in their evaluations. Further differences include the fact that ECIS and ER featured a criterion on the *significance* of the research, while BPM 2007 featured a criterion *practical impact*. In Table 1, we summarize the review criteria used across the three conferences, show which of the criteria are similar across the conferences, and also indicate which review criterion addresses aspects of *relevance* or *rigour*, or address *other* characteristics of a submission.

Visual inspection of Table 1 leads to a number of preliminary conclusions. Most notably, the set of criteria in use for different IS conferences appear to differ considerably. For instance, while ECIS 2007 featured two criteria that could nominally be associated with 'rigour' (i.e., theoretical strength and methodology used), BPM 2007 and ER 2007 featured only one such criterion (technical soundness and technical quality, respectively). Second, while all three conferences featured a dedicated 'relevance' criterion, in all cases did it address the submission fit with the theme of the conference or the track. This understanding of relevance is considerably different from the prevalent understanding of *practical* research relevance, i.e., the extent to which scholarly work is interesting and important, applicable, current and accessible to practitioners (Benbasat and Zmud 1999;

Klein, Jiang and Saunders 2006). In fact, only one of these three conferences (BPM 2007) had a criterion dedicated to relevance to practitioners (practical impact). ECIS 2007 featured a criterion 'significance/contribution' and ER 2007 featured a criterion 'significance', both of which, one may argue, could be seen as addressing the aspect of 'importance' of relevant research as defined by Klein et al. (2006). Similarly, BPM 2007 and ER 2007 evaluated papers in terms of their 'originality' – which, again we speculate, could address an aspect of practical relevance. Third, all conferences employ two or three criteria that we find address criteria pertaining only the organization of the conference – the relevance to the track theme, the clarity of the presentation of the research ('presentation'), and the 'appeal to audience' (at ECIS 2007).

Table 1: Conference Review Criteria

ECIS 2007	**BPM 2007**	**ER 2007**	**Criterion nature**
	Technical Soundness	Technical Quality	*Rigour*
Theoretical Strength (the strength of the theoretical foundations used, if any)			
Methodology Used (the quality of the methodology and/or analytical techniques in use)			
	Practical Impact		*Relevance*
Significance/Contribution (the likely significance and potential contribution to the field)		Significanc e	
	Originality	Originality	
Relevance to ECIS (the submission fit with the theme of the conference and the track)	Relevance to BPM	Relevance to ER	*Other*
Presentation (the clarity of organization, the presentation, and the writing)	Presentatio n	Presentation	
Appeal to Audience (the likelihood of the paper drawing and keeping an audience at the conference)			

These findings lead to one interesting question that requires further discussion and testing: Can it be guaranteed that our conference review systems assess a submission's rigour and relevance – or do the review criteria commonly employed at IS conferences actually capture some other aspects of a submission? Most of the commonly used evaluation criteria seem not to reflect the popular rigour vs. relevance discussion. Criteria such as presentation, or appeal to audience do not appear to map to rigour or relevance in a straight forward manner.

DATA ANALYSIS

What is the actual relationship between rigour and relevance in conference submissions?

In order to test the first research proposition (rigour and relevance very much can complement each other, and are actually not in a contradictory relationship to each other in a conference submission), we first had to estimate a proxy for rigour and relevance, for all conference papers across the three conferences considered. To that end, we computed for each submission the average total factor score for those review criteria that we were able to map to either rigour, or relevance, as per Table 1 above. Then, we normalised both the rigour and relevance scores to all scale from -1 to 1, so that a value of 0 represents the medium score on the Likert scale that has been used, i.e. for example a 4 in a 7-point Likert scale. This approach allowed us to make the three conference sets, and the different sets of review criteria used, comparable.

Starting with the conference from our set of three conferences that had the highest number of submissions, the ECIS conference, we derive the diagram depicted in Figure 3. The two different colours in this figure reflect the pool of accepted and of rejected papers.

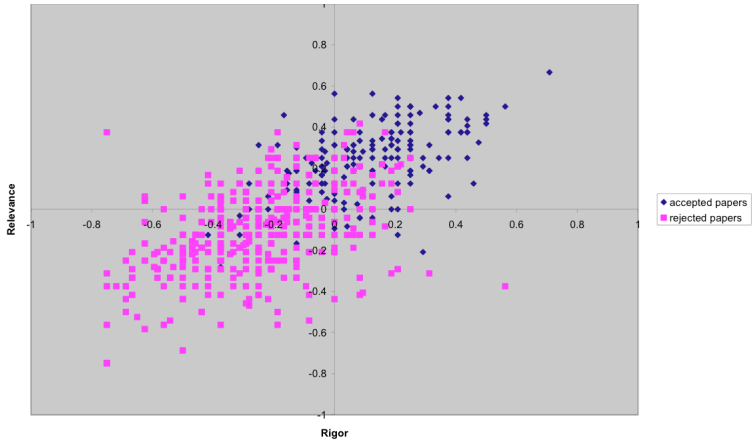

Figure 3: The distribution of papers at the ECIS 2007 conference

Visual inspection of Figure 3 confirms that the distribution of the conference paper submissions forms a shape that is much closer to the left case in Figure 1 than the right case. In fact, aside from a small number of (ultimately rejected) papers on the far top left and the far bottom right, there are no strong signs for papers that exclusively emphasis either rigour or relevance. Similar diagrams are shown in Figure 4 below for the ER and the BPM conference, both of which received significantly less submissions, and also each have a dedicate conference focus (conceptual modelling, and business process management, respectively).

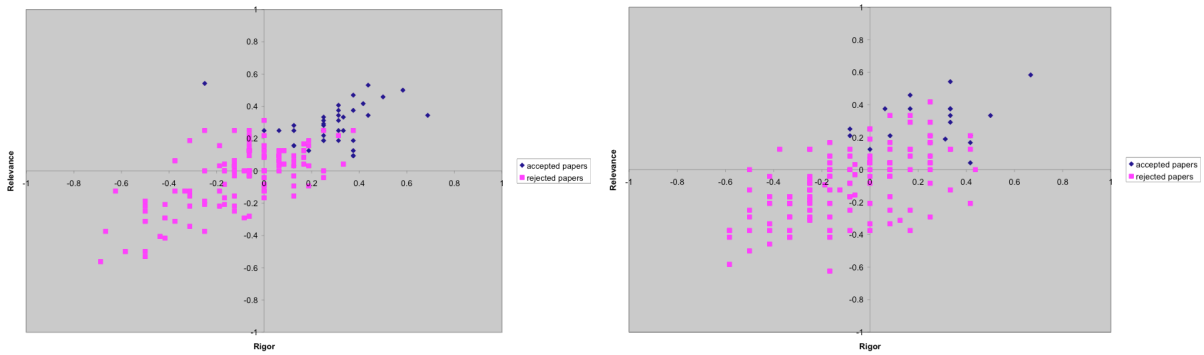

Figure 4: The distribution of papers at the BPM 2007 (left) and the ER 2007 conference (right)

Figure 4 shows that even for the much smaller and more focused conferences ER 2007 and BPM 2007, it can be seen that the papers are distributed roughly along the 45 degree curve indicating again that rigour and relevance are indeed highly complementary to each other in papers submitted to typical IS conferences.

All three data analyses point towards a strong confirmation of our first research proposition, i.e. there is no indication that a rigour *versus* relationship exists within papers produced by the IS community. The next logical question, consequently, is then which of these papers is ultimately accepted at IS venues? We examine our corresponding research proposition P2 below.

Do IS conference acceptance decisions maintain the balance between rigour and relevance?

In order to examine our research proposition P2 – acceptance decisions made at IS conferences are skewed towards rigour rather than relevance – we performed a second test on the data set obtained. Similar to the analysis done for all conference submissions, we computed normalized rigour and relevance scores for accepted papers and rejected papers separately. The corresponding scores are provided in Table 2. Based on the data displayed in Table 2, and to achieve a visualisation that corresponds with Figure 2, we the plotted the average scores of all submissions, and of the accepted papers only, at the three conference considered in Figure 5.

Figure 5 yields interesting insights. First of all, it visualizes the distance (both on the x- and the y-axis) between all conference submissions and accepted submissions. This informs us about the quality standards of the respective conference, which is reflected, for instance, in the acceptance rate. ECIS 2007 has the highest acceptance rate (34.5 per cent) – visualized in a shorter distance between all submissions and accepted submissions. In contrast, BPM 2007 accepted only 13.8 per cent – which could mean that the quality standards (in terms of rigour and relevance) of an accepted paper are much higher than those of all paper submissions

(note, however, that it is not necessarily true that the claim for higher-quality is associated with lower acceptance rates. It may also be that certain conferences attract higher-quality papers than others).

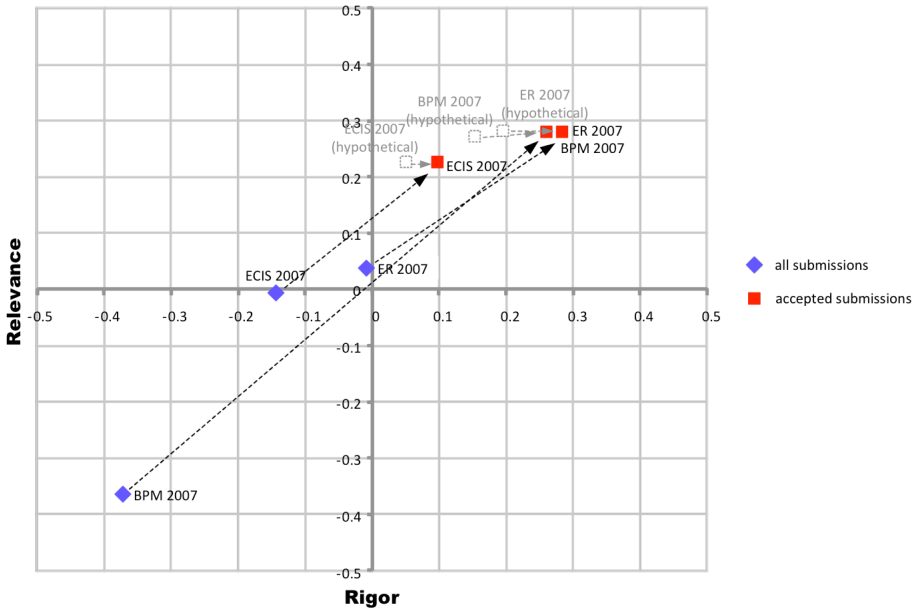

Figure 5: The deviation of rigour and relevance of accepted papers in contrast to all paper submissions

Figure 5 further informs us about how far acceptance decisions made by the conference editorial teams deviate from a 'perfect' balance between rigour and relevance. Recall, we speculated in scenario II of Figure 2 above that a perfect balance between rigour and relevance between accepted papers would be achieved if accepted papers would be placed in a rigour/relevance scatter plot on a 45 degree straight line moved further to the top right from all submissions, which would indicate that, per conference, accepted papers would simply achieve better scores (for both rigour and relevance) in comparison to all submissions. We indicated these hypothetical scatter points as grey dashed boxes in Figure 5, and also indicate the true placement of the average scores for accepted papers (the red scatter points).What is notable from Figure 5 as well as the data displayed in Table 3 is that accepted papers are, by trend, skewed towards an additional emphasis on rigour. This is indicated by the right-shift of the scores from hypothetical accepted papers towards actually accepted papers. We also realize that this shift is stronger for the smaller, more focuses conferences (ER and BPM 2007) than for ECIS 2007. Overall, in light of the data obtained, our analysis and visualization indicates that conference editorial teams are skewed towards rigour in their acceptance decisions, which suggests to tentatively accept proposition P2.

Table 2: Rigour and Relevance Scores

Rigour scores						
	all		*accepted*		*rejected*	
	N	Mean	N	Mean	N	Mean
ECIS 2007	580	-0.1407	200	0.0981	380	-0.2664
BPM 2007	152	-0.3717	21	0.2599	131	-0.0851
ER 2007	160	-0.0076	38	0.2791	122	-0.0968
Relevance scores						
	all		*accepted*		*rejected*	
	N	Mean	N	Mean	N	Mean
ECIS 2007	580	-0.0028	200	0.2256	380	-0.1230
BPM 2007	152	-0.3608	*21*	0.2808	131	-0.0872
ER 2007	160	0.0389	*38*	0.2804	122	-0.0364

Our data also allows us to calculate the actual impact of conference editors on the acceptance decision in a Rigour/Relevance metric. In order to do this, we calculated the following values for each of the three conferences and a consolidated non-weighted value across all three conferences (*Rigour/Relevance metric* = average score for rigour of accepted papers / average score for relevance of accepted papers). A value of more

than 1 means that the conference editors put more emphasis on rigour than relevance, while a value of less than 1 would mean that relevance plays a more important role. Table 3 gives the results obtained.

Table 3: Rigour/Relevance Metric

	accepted	**rejected**
ECIS 2007	0.4348	2.1659
BPM 2007	0.9256	0.9759
ER 2007	0.9954	2.6593

We see from Table 3 that the ER 2007 conference was closest to obtaining a good balance between rigour and relevance for their accepted papers, while across all rejected papers at ER 2007, the scores indicate an over-emphasis of relevance (a metric score of 2.66). BPM 2007 is also close to such a balance, with a metric score of 0.93 for accepted papers. Interestingly, also across all rejected papers can we observe the balance between rigour and relevance (a score of 0.98). In stark contrast, the ECIS 2007 conference achieved a metric of 0.43, indicated that papers are attributed by higher scores on relevance metrics than on rigour metrics. Overall, the data obtained in Table 3 corroborates our findings discussed above.

CONCLUSIONS

This paper makes a key contribution to the rigour versus relevance discussion by sharing empirical evidence on the actual relationship between rigour and relevance in acceptance decisions made in IS conference reviewing processes. Based on data from three recent prominent IS conferences, we were able to better understand the actual significance of criteria as they are used within the review process.

The most important findings are twofold. First, rigour and relevance do not have to compromise each other. A detailed study of the characteristics of all papers submitted to these three conferences showed that the typical submission to an IS conference actually balances rigour and relevance very well. Second, we established findings that suggest that editorial decisions tend to emphasis – to a certain extent – rigour over relevance in their acceptance decisions, leading to published papers that are more characterized by rigour than relevance.

The presented research findings have to be contextualised in light of some limitations. First, the source of empirical evidence is limited to only three conferences. It also does not comprise any longitudinal analyses. Second, in order to make the conferences comparable, we had to map selected review criteria to a normalized rigour and relevance score. While this mapping exercise was performed to the satisfaction of each member of the research team, we cannot eliminate all potential threats of mapping bias to our findings, and the implications thereof. Third, the data set obtained may also be subject to some bias. Specifically, ECIS 2007 had the conference theme “Relevant Rigour – Rigorous Relevance” – which, in turn, may have contributed to ECIS 2007 being unusually sensitive to rigour or relevance review criteria. Fourth, we considered review scores as a primary data source. While this data gives us insights into the reviewer’s perceptions of rigour and relevance of a manuscript, the ultimate decision about acceptance is made by the conference editors and track chairs. The difference between peer assessments and editor assessments may cause a bias that we did not considered in great detail.

While these limitations certainly denote key boundaries to the generalizability of our findings, we nevertheless deem them to be a strong indicator of IS review practices overall. We would expect the tenet of our analysis to hold for other IS conferences also. Moreover, we feel that our research has the potential to trigger a variety of related research streams:

First, we encountered a variety of evaluation criteria, and obviously lack as a discipline a core set of values when it comes to the set of evaluation criteria. We see potential to propose at least a superset of well-defined and clearly categorised evaluation criteria that could be used to pre-populate conference review systems. The work reported by Recker et al. (2009) towards an IS relevance index is an example for the desire to build a set of criteria reflecting facets of publication quality to relevance. It will be important, however, to complement this work with relevance judgements from outside the academic IS community, and to carefully consider contrasting viewpoint about such review standardization efforts.

Second, a better understanding of the relative influence of various review criteria would allow more reliable cross-conference comparisons. We see the proposed rigour/relevance metric as a potential important indicator that could be published as part of the conference proceedings. This would allow scholars and practitioners alike to reflect upon the nature of a conference and the type of papers that can be expected at these conferences prior to deciding to submit and/or attend.

Third and finally, we see the potential to further extend our simple causal model by a deeper analysis of the antecedents of IS relevance and IS rigour. Focus groups, Delphi studies and surveys of IS academics could be

adequate research methods to gain further insights what actually constitutes rigour or relevance. This would allow providing more normative, pro-active guidance to authors of papers as well as to reviewers and editors. We see our work a starting point in this endeavour and hope that other scholars join in this challenge.

REFERENCES

Agarwal, R & Lucas Jr., HC 2005, 'The Information Systems Identity Crisis: Focusing on High-Visibility and High-Impact Research', *MIS Quarterly,* vol. 29, no. 3, pp. 381-398.

Avgerou, C, Siemer-Matravers, J & Bjørn-Anderson, N 1999, 'The Academic Field of Information Systems in Europe', *European Journal of Information Systems,* vol. 8, no. 2, pp. 136-153.

Baskerville, R 2007, 'Forms of Design Research and the Role of Rigor and Relevance', *Wirtschaftsinformatik,* vol. 49, no. 5, pp. 403-404.

Baskerville, R & Myers, MD 2002, 'Information Systems as a Reference Discipline', *MIS Quarterly,* vol. 26, no. 1, pp. 1-14.

Benbasat, I & Weber, R 1996, 'Research Commentary: Rethinking "Diversity" in Information Systems Research', *Information Systems Research,* vol. 7, no. 4, pp. 389-399.

Benbasat, I & Zmud, RW 1999, 'Empirical Research in Information Systems. The Practice of Relevance', *MIS Quarterly,* vol. 23, no. 1, pp. 3-16.

Darroch, F & Toleman, M 2005, 'Bridging the Academia-Industry Divide: Academics Reach Out!', *Software Practitioner,* vol. 15, no. 6, pp. 3-7.

Davenport, TH & Markus, ML 1999, 'Rigor vs. Relevance Revisited: Response to Benbasat and Zmud', *MIS Quarterly,* vol. 23, no. 1, pp. 19-23.

Dennis, AR, Valacich, JS, Fuller, MA & Schneider, C 2006, 'Research Standards for Promotion and Tenure in Information Systems', *MIS Quarterly,* vol. 30, no. 1, pp. 1-12.

Dubé, L & Paré, G 2003, 'Rigor in Information Systems Positivist Case Research: Current Practices, Trends, and Recommendations', *MIS Quarterly,* vol. 27, no. 4, pp. 597-635.

Grover, V, Lee, CC & Durand, D 1993, 'Analyzing Methodological Rigor of MIS Survey Research from 1980-1989', *Information & Management,* vol. 24, no. 6, pp. 305-317.

Klein, G, Jiang, JJ & Saunders, C 2006, 'Leading the Horse to Water', *Communications of the Association for Information Systems,* vol. 18, no. 13, pp. 259-274.

Recker, J, Young, R, Darroch, F, Marshall, P & McKay, J 2009, 'ACIS 2007 Panel Report: Lack of Relevance in IS Research', *Communications of the Association for Information Systems,* vol. 24, no. 18, pp. 303-314.

Robey, D & Markus, ML 1998, 'Beyond Rigor and Relevance: Producing Consumable Research about Information Systems', *Information Resource Management Journal,* vol. 11, no. 1, pp. 7-15.

Rosemann, M 2008, 'Understanding and Impacting the Practice of Business Process Management' in Dumas, M, Reichert, M & Shan, M-C (eds.) *Business Process Management - BPM 2008*, Springer, Milan, Italy, p. 2.

Rosemann, M & Vessey, I 2008, 'Toward Improving the Relevance of IS Research to Practice: The Role of Applicability Checks', *MIS Quarterly,* vol. 32, no. 1, pp. 1-22.

Senn, JA 1998, 'The Challenge of Relating IS Research to Practice', *Information Resource Management Journal,* vol. 11, no. 1, pp. 23-28.

Straub, DW, Boudreau, M-C & Gefen, D 2004, 'Validation Guidelines for IS Positivist Research', *Communications of the Association for Information Systems,* vol. 13, no. 24, pp. 380-427.

Wand, Y & Weber, R 2002, 'Research Commentary: Information Systems and Conceptual Modeling - A Research Agenda', *Information Systems Research,* vol. 13, no. 4, pp. 363-376.

Whitley, EA & Galliers, R 2007, 'An Alternative Perspective on Citation Classics: Evidence from the First Ten Years of the European Conference on Information Systems', *Information & Management,* vol. 44, no. 5, pp. 441-455.

ACKNOWLEDGEMENTS

We wish to thank the organizing committee of ECIS 2007, BPM 2007 and ER 2007 for their willingness t contribute to this study.

COPYRIGHT